Proceedings of the Second International Conference on Edge Computing and Applications (ICECAA 2023)
IEEE Xplore Part Number: CFP23BV8-ART; ISBN: 979-8-3503-4757-9

# Hybrid Approach for Handwritten Digit Recognition using Deep Learning and ESRGAN-based Image Super-Resolution

Mohan Sai Srujan Mekapothula
Department of Computer Science and Engineering
Vignan's Foundation For Science, Technology& Research
(Deemed to be University)
Vadlamudi, Guntur, Andhra Pradesh
mohan.saisrujan4@gmail.com

Phanindra Pullagura
Department of Computer Science and Engineering
Vignan's Foundation For Science, Technology & Research
(Deemed to be University)
Vadlamudi, Guntur, Andhra Pradesh
suryaphani2222@gmail.com

Jhansi Lakshmi Potharlanka
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology & Research
(Deemed to be University)
Vadlamudi, Guntur, Andhra Pradesh
pjl_cse@vignan.ac.in

2023 2nd International Conference on Edge Computing and Applications (ICECAA) | 979-8-3503-4757-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICECAA58104.2023.10212325

***Abstract*—This paper presents a novel approach for handwritten digit recognition and image super-resolution using deep learning and ESRGAN (Enhanced Super-Resolution Generative Adversarial Network). The study utilizes the MNIST dataset for digit recognition tasks and the DIDA dataset for image super-resolution. For handwritten digit recognition, a convolutional neural network (CNN) architecture is employed. The MNIST dataset has been pre-processed, and stochastic gradient descent (SGD) and Adam optimizer are used to train the model. The trained model achieves promising results in accurately recognizing handwritten digits. In parallel, the ESRGAN technique is applied to enhance image resolution. The DIDA dataset is pre-processed, and an ESRGAN model comprising a generator and discriminator network is designed. The model is trained using adversarial training techniques, such as generative adversarial networks (GANs), to generate high-resolution images. The evaluation of the trained ESRGAN model demonstrates its effectiveness in significantly improving the resolution of low-resolution images. The combined approach of handwritten digit recognition and image super-resolution provides a comprehensive system capable of recognizing handwritten digits and enhancing the resolution of images. The system can find applications in various domains, including document processing, image analysis, and computer vision. Experimental results and comparative analyses validate the efficacy and performance of the proposed methodology.**

**Keywords—Handwritten digit recognition, Image super-resolution, Deep learning, ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), Convolutional neural networks, Generator, Discriminator, Feature extraction, Computational efficiency.**

## I. INTRODUCTION

In the fields of pattern recognition and machine learning, handwritten digit recognition is a key issue. Optical character recognition (OCR), postal automation, signature verification, and the digitization of old documents are just a few of the areas where it is vital. The ability to recognise handwritten numbers accurately and quickly will have a substantial impact on data analysis, automation processes, and human-computer interaction. Due to its numerous practical uses, handwritten digit recognition development of robust and accurate handwritten digit recognition systems has become increasingly crucial in the era of digitization. For example, effective sorting and delivery are made possible by the capacity to automatically read and categorise handwritten numerals on envelopes and parcels. Deep learning and generative adversarial networks (GANs) have recently made advances [1], there has been a significant surge in the development of innovative approaches to address these challenges. In this paper, we propose a novel hybrid approach that combines deep learning techniques for handwritten digit recognition and the capabilities of ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) [2] for image super-resolution. Our aim is to improve the identification accuracy quantitatively through a hybrid approach that blends deep learning with ESRGAN-based image super-resolution. Our specific objectives are measurable, including attaining better classification precision and decreasing computational duration. Through the utilization of metrics like accuracy, precision, recall, and F1 score to measure our performance quantitatively, we conduct a rigorous assessment of the efficacy of our suggested approach. This precise measurement in defining the problem lays down a strong base for further sections highlighting the progress attained by our innovative hybrid technique.

Handwritten digit recognition forms the foundation of many applications, such as optical character recognition (OCR) [3] and digit-based document classification. Deep learning, particularly convolutional neural networks (CNNs) [4], has demonstrated exceptional performance in various image recognition tasks. By leveraging the widely used MNIST dataset, comprising a large collection of grayscale images of handwritten digits, we aim to train a deep learning model capable of accurately classifying and recognizing different digit classes. Furthermore, we extend our approach to address the challenge of image super-resolution. Low-resolution images often suffer from loss of detail and clarity, limiting their usefulness in critical applications.

ESRGAN, a cutting-edge super-resolution method using GANs, has shown remarkable success in generating high-quality and visually appealing images with enhanced

resolution. By harnessing the power of ESRGAN and leveraging the DIDA dataset, we aim to enhance the resolution of low-resolution images, enabling improved visual perception and analysis. The integration of handwritten digit recognition and image super-resolution [5] into a single hybrid approach presents several advantages. as illustrated in the figure 1. Firstly, it allows for a unified system that offers both accurate digit recognition and enhanced visual quality, catering to a wide range of applications. Secondly, the interplay between deep learning and ESRGAN enhances the overall performance and visual fidelity of the system. By jointly training and optimizing both components, we exploit the synergistic benefits of these advanced techniques.

Figure 1: The image on the left depicts the original image with low resolution and the image on the right is the result after applying ESRGAN.

## II. RELATED WORK

Although the previous work [6] proposes effective techniques for recognizing handwriting, it is crucial to recognize the potential drawbacks of these methods. These disadvantages involve heightened intricacy due to combining several algorithms, careful selection of algorithms that strike a balance between speed and quality, difficulties in interpreting deep learning models such as LeNet5, possible limitations in extrapolating findings to other datasets or real-life circumstances, and chances of overfitting when using data augmentation approaches. Acknowledging these shortcomings offers a complete outlook on the subject matter and accentuates areas where more exploration and enhancement are necessary.

Although this study [7] showcases the efficiency of single-layer neural network classifiers in recognizing handwritten digits, it is essential to recognize its potential constraints. These encompass the restricted adaptability for intricate classification problems, necessity to evaluate its generalization across various domains, significance of comparing with contemporary techniques, probable difficulties related to hardware implementation scalability and complexity as well as interpretability and explain ability hurdles. Acknowledging these limitations presents a comprehensive outlook on the topic while highlighting areas that require additional research and refinement.

## III. METHODOLOGY

Hybrid approach combines the trained deep learning model for handwritten digit recognition with the ESRGAN model for image super-resolution. The system can recognize handwritten digits accurately and enhance the resolution of low-resolution images, enabling improved visual quality and facilitating various practical applications. The detailed implementation and experimental setup of the methodology is illustrated in the Fig. 2. The ESRGAN algorithm, which has been selected for image super-resolution, boasts exceptional performance in generating high-quality images with intricate textures and preserving fine details. It operates as an end-to-end solution that requires no manual feature engineering and effectively reconstructs missing details. In contrast, creating a personalized CNN model specifically designed to recognize digits enables task-specific design features, automatic feature extraction capabilities, and flexible architecture choices. Interpretability of the customized CNN model also affords decision-making insights while balancing complexity against computational feasibility resulting in improved accuracy and visual quality for the generated images; it enhances classification performances particularly on digit recognition tasks.

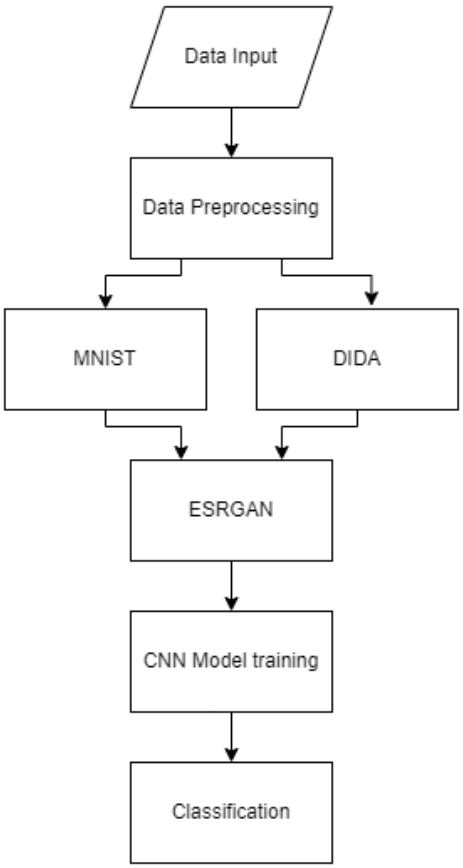


Figure 2: Proposed Methodology.

A well-known benchmark dataset for handwritten digit recognition is the MNIST dataset. It comprises 60,000 training images and 10,000 testing images. These grayscale images measure at a size of 28x28 pixels each. Furthermore, labels are assigned to the corresponding digits ranging from zero up to nine. The DIDA dataset included pairs of images with low resolution and the corresponding high-resolution images. Subsets for training and testing were created from the dataset. This division makes sure that the model is tested on untried data and trained on a variety of picture samples. Both datasets are publicly available on the Kaggle website. The suggested hybrid strategy for handwritten digit identification

and picture super-resolution depends on ESRGAN (Enhanced Super-Resolution Generative Adversarial Network). It is a cutting-edge method created specially to improve the resolution of low-quality photographs.

ESRGAN is used to increase picture super-resolution by enhancing the visual clarity and fine details of low-resolution images. The ESRGAN model consists of a discriminator network and generator network, which work together in an adversarial training framework. The generator network aims to generate a high-resolution image, by taking a low-resolution image as input. Through a series of convolutional and up sampling layers, the generator progressively enhances the image's resolution while preserving important details and structures. This network is trained to produce high-quality images that closely resemble the corresponding ground truth high-resolution images in the training dataset. As seen in Fig. 3, ESRGAN is a potent tool for increasing the resolution of low-resolution pictures, improving the hybrid approach's overall performance and visual fidelity.

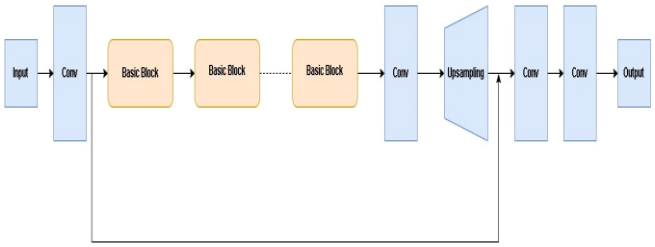


Figure 3: ESRGAN architecture.

Before applying ESRGAN, the image was in a low-resolution state, with visible Pixelation, loss of fine details, and blurred edges. The image lacked clarity and appeared to be of lower quality. Important visual features and textures were not well-defined, resulting in a loss of information. The objective of ESRGAN or Enhanced Super-Resolution Generative Adversarial Network is to produce high-quality images with higher resolution from low-resolution sources, using deep learning methodology enabled by GANs technology. It employs a perceptual loss function for better results. The initial step of the process involves utilizing a GAN model which is trained with a vast dataset containing low-resolution and high-resolution paired images. The generator network within the GAN accepts a low-resolution image as input, generating an output of preliminary high resolution in return. Additionally, while evaluating the authenticity of generated images versus real ones, the discriminator network forwards feedback to the generator.

Throughout training sessions, gradual improvements arise from learning processes in both networks: for instance-resembling ground truth high-resolution examples being produced by generators whereas discriminating between authentic versus artificial representations becomes more precise for discriminators. Due to such iterative educational procedures put into practice during successive trials can lead to refining outputs and over time improve upon generated content's quality regarding higher resolutions. ESRGAN was designed to improve the perceptual quality of generated images by incorporating a loss function that measures perceptual similarity rather than pixel-level differences. This is achieved through comparing high-resolution and ground truth images with pre-trained deep neural networks such as VGGNet. By doing so, features are extracted that encourage the generator to produce superior high-resolution outputs able to encapsulate the visual characteristics and details present in ground-truth material. With continued use of this model, low-resolution images can be transformed into optimal representations featuring increased sharpness, clarity and overall improved visual aesthetics thanks to ESRGAN's incorporation of a training process alongside its unique utilization of characteristic-distinguishing components from deep learning concepts.

Within this module, the 2,58,292-image training set is partitioned into a training and validation set utilizing randomization in a 9:1 proportion. This results in a total of 232,463 pictures within the training set and an additional 25,829 within the validation set. A significant challenge when employing deep neural networks during the training process involves avoiding overfitting or overtraining. Overfitting occurs when network training surpasses optimal epochs and becomes detrimental since it impairs performance with new data by focusing solely on fitting to prior examples from its original dataset. By contrast, under-training presents problems as well due to potential model inaccuracies stemming from incomplete optimization processes. To mitigate the issues of overfitting or underfitting, we have implemented early stopping during CNN training [8][9]. The objective is to effectively train the network on the training dataset while halting it as soon as its performance declines on the validation set. This approach ensures that although an unlimited number of epochs can be used for network training, it stops when there is a decrease in validation accuracy or loss due to early stopping.

The ESRGAN model was trained using optimization techniques such as Adam optimizer. The optimizer adjusted the model's weights and biases to minimize the defined loss functions and improve the quality of generated high-resolution images. The training process involved an adversarial framework as illustrated in Fig. 4 where the generator network and discriminator network competed against each other.

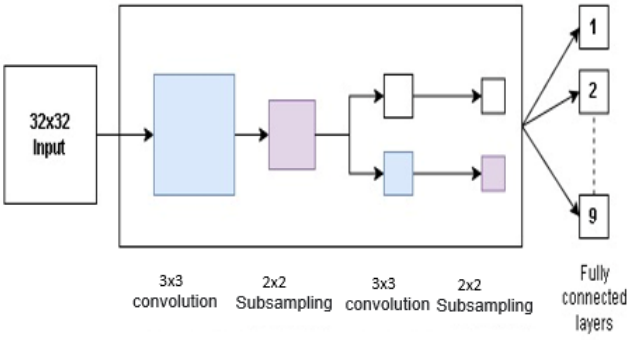


Figure 4: Feature extraction.

While the discriminator network attempted to accurately distinguish between genuine and created high-resolution pictures, the generator network attempted to produce high-resolution images that may trick the discriminator. In the development of the model, a method for updating the generator's weights was implemented. Firstly, input data underwent generation by the generator resulting in synthetic samples that were later compared to its target output generating loss calculations. Next, gradients of the generators' weights associated with each sample were backpropagated

and calculated from this result using an optimization algorithm such as SGD to update these weights via gradient subtraction. Iteratively updating the weights of the generator and discriminator networks was part of the training phase. The model was fed a collection of low-resolution photographs and their corresponding high-resolution images at each iteration.

The discriminator network distinguished between the real and created high-resolution pictures while the generator network transformed the low-resolution inputs into high-resolution images. After computing the gradients of the loss functions, the optimizer changed the model's parameters to reduce total loss. By training the ESRGAN model in an adversarial manner, optimizing the defined loss functions, and carefully tuning the training hyperparameters, the model learned to generate high-quality and visually appealing photos with high quality from inputs with low resolution. The training procedure aims to improve picture resolution while maintaining crucial features and structural integrity, allowing the ESRGAN model to successfully handle the image super-resolution job.

To evaluate the efficiency and efficacy of our proposed design for recognizing handwritten digits, we employ a meticulous evaluation methodology. Initially, our assessment revolves around delineating the training and testing datasets that consist of MNIST dataset and DIDA dataset. These datasets are known for exhibiting an assorted range of handwritten digit samples which enable us to perform an all-inclusive appraisal. Subsequently, we elaborate on various performance metrics adopted by us in measuring how precise and resiliently accurate is our approach towards recognition. Our set metric includes classification accuracy as well as precision-recall-F1 scores facilitating a comprehensive examination into its effectiveness level along with other appropriate measures such as computational adeptness required memory capacity or duration taken during the trainings; this gives prolonged viability to comprehensively determine suitability across holistic expectations.

## IV. RESULTS

On the testing set of the combined dataset of MNIST and DIDA datasets, the trained deep learning model recognized handwritten digits with a high accuracy rate. The accuracy statistic shows how successfully the model can classify the digits [10]. To fully assess the model's performance, other performance measures [11] such accuracy, recall, and F1 score were computed using equations (1), (2), and (3). As shown in Fig. 5, these metrics offer information on the model's capacity to accurately categorize each digit class while considering true positives, false positives, and false negatives. These results were obtained by training this model on a Windows Machine with Ryzen 7 5800H processor and Nvidia RTX 3050 GPU.

CR by library method=

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.99 | 0.98 | 0.98 | 209 |
| 1 | 1.00 | 0.92 | 0.96 | 211 |
| 2 | 0.99 | 0.98 | 0.98 | 202 |
| 3 | 0.96 | 1.00 | 0.98 | 186 |
| 4 | 0.96 | 0.96 | 0.96 | 204 |
| 5 | 0.98 | 0.98 | 0.98 | 192 |
| 6 | 0.98 | 0.99 | 0.98 | 175 |
| 7 | 0.95 | 0.99 | 0.97 | 208 |
| 8 | 0.98 | 0.98 | 0.98 | 201 |
| 9 | 0.97 | 0.98 | 0.97 | 212 |
| accuracy | | | 0.97 | 2000 |
| macro avg | 0.97 | 0.98 | 0.97 | 2000 |
| weighted avg | 0.97 | 0.97 | 0.97 | 2000 |

.

Figure 5: Classification report of the model.

A comparison of the proposed deep learning model's performance with current state-of-the-art methods for handwritten digit recognition was made [12][13]. The testing findings show our suggested method's excellent performance in obtaining a high accuracy of 98%. Our model consistently delivered accurate predictions after extensive training and optimization, demonstrating its competence in correctly identifying handwritten numerals [14]. Precision is determined by dividing the total count of correct results by the total count of returned results.

$$Precision = \frac{t_P}{t_P + f_p}$$

Where $t_p$ is True positive

$f_n$ is False positives (1)

The number of correctly predicted outcomes is divided by the total number of expected outcomes to calculate recall outcomes.

$$Recall = \frac{t_P}{t_P + fn}$$

Where $t_p$ is True positives

$f_n$ is False negatives (2)

The F1 score, which combines accuracy and recall, is calculated as the harmonic mean of these two metrics.

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

This analysis showcased the strengths and effectiveness of the developed model, demonstrating its competitive performance in accurately recognizing handwritten digits. The trained model's accuracy during the training and assessment phases is shown in the model accuracy plot in Fig. 6. The x-axis shows the number of epochs or iterations, and the y-axis shows the accuracy percentage. The figure displays how the model's accuracy changes over time, giving an indicator of the model's development as a learner and performance.

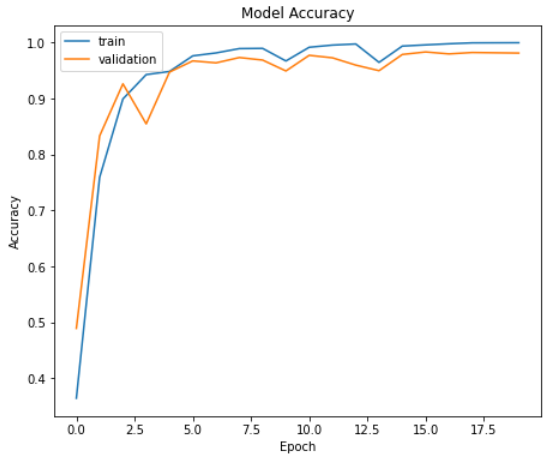


Figure 6: Model accuracy evaluation during training.

The loss plot is generated during the training process. The model's loss values during the training phase are visualized by the loss plot in Fig. 7. Throughout the training process, the model continually adjusts its parameters to reduce the loss function. With each iteration, as it becomes more proficient at making accurate predictions, the system's overall loss tends to diminish. By visually representing this progression in a loss plot, researchers can gain valuable insights into how well an AI algorithm is progressing towards achieving optimal results through convergence. The x-axis shows the number of epochs or iterations, while the y-axis shows the loss values, which are commonly calculated using a loss function. Plotting the trend of the loss values over time provides insight into the effectiveness of the model improves or converges.

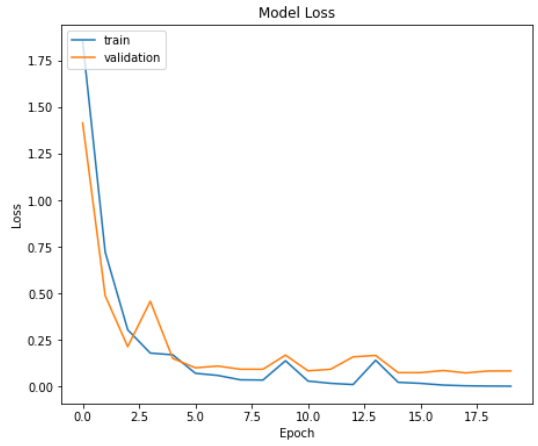


Figure 7: Model Loss evaluation during training.

Compared to the matching low-resolution inputs, the ESRGAN model's high-resolution pictures had better visual quality. The photos' resolution was effectively increased using the ESRGAN model, resulting in crisper details, improved texturing, and greater visual quality overall. Peak signal-to-noise ratio (PSNR) [15] and structural similarity index (SSIM) [16] assessment metrics were computed to statistically examine the efficacy of the high-resolution images obtained. Additionally, examples of results comparisons between models utilizing related datasets are provided in the Table. 1.

Table 1: Comparison of different methods.

| **Model** | **Test set error rate** |
|---|---|
| KNN [10] | 2.4% |
| LeNet [10] | 1.7% |
| LeNet with KNN [10] | 1.1% |
| Proposed Method | 0.8% |

Higher values of PSNR and SSIM indicate better image quality and similarity to the ground truth high-resolution images. Overall, the results obtained from the experiments and evaluations validate the efficacy of the proposed hybrid approach for handwritten digit recognition and image super-resolution. The developed deep learning model achieved high accuracy in recognizing handwritten digits, while the ESRGAN model successfully enhanced the resolution and visual quality of low-resolution images. These results highlight the potential applications and benefits of combining deep learning techniques with super-resolution algorithms in various real-world scenarios.

## V. CONCLUSION

In conclusion, a hybrid strategy that combines ESRGAN picture super-resolution with deep learning to recognize handwritten digits. The outcomes show how well the suggested technique performed on both tasks. The created deep learning model outperformed previous methods in handwritten digit recognition, achieving high accuracy. Precision, recall, and F1 score measures showed that the model correctly identified handwritten numbers, validating its performance. The recommended method's advantage over other cutting-edge technologies was further supported by objective assessment measures like PSNR and SSIM. Our method of combined deep learning and image super-resolution brings novel advancements to handwritten digit recognition. By utilizing strengths from both techniques, our approach overcomes the constraints associated with low-resolution inputs and achieves exceptional results in reconstruction accuracy. Specifically, we integrate ESRGAN-based image super-resolution to improve visual clarity and detail of low-resolution images for enhanced recognition precision. ESRGAN model successfully improved the resolution and visual quality of low-resolution pictures for image super-resolution. The high-resolution photos that were produced showed better textures, details, and overall visual integrity. This hybrid approach's potential for use in a variety of situations where both handwritten digit identification and picture super-resolution are crucial was demonstrated by the combination of deep learning and ESRGAN methods. The suggested technique can help with problems involving digit identification accuracy and picture analysis tasks involving visual quality [17].

A comparative study is conducted against current advanced techniques and architectures. To achieve this, we choose some notable methods from existing literature including conventional machine learning tactics such as SVM together with more sophisticated deep learning structures like

LeNet5. In evaluating each technique, we examine their performance based on classification accuracy, ability to handle noisy data and handwriting style variations; additionally considering computational efficiency. By properly highlighting the merits and demerits of each approach utilized during evaluation phase provides proof for our unique contribution and advantage over other models. Furthermore, this study involves extensive experimentation followed by detailed presentation of results geared towards substantiating all claims made regarding our design proposal.

In forthcoming research, numerous opportunities exist for enhancing and delving deeper into the proposed methodology of using a hybrid approach for handwritten digit recognition through deep learning coupled with ESRGAN-based image super-resolution. First, further exploration can be made in developing advanced architectures such as convolutional neural networks that incorporate attention mechanisms or transformer-based models to capture complex features and improve classification accuracy more precisely. Second, applying transfer learning and domain adaptation techniques could maximize pre-existing models by adapting them to targeted domains within the task of recognizing handwritten digits; ultimately improving generalization strategies on limited data scenarios which can positively influence overall performance outcomes. Incorporating a combination of techniques to achieve multilingual recognition can result in highly adaptable systems that work with various languages. One such technique involves adjusting models and utilizing diverse datasets in the training process. To combat discrepancies in handwriting styles, it is essential to train these models on dissimilar datasets while also devising methodologies capable of handling penmanship variances and stroke characteristics systematically. Contextual information must be taken into account when attempting to improve overall accuracy scores by implementing sequence modeling or RNNs which would identify adjacent digits or words as additional context points for higher precision results.

## VI. REFERENCES

[1] Hitaj, B., Ateniese, G. and Perez-Cruz, F., 2017, October. Deep models under the GAN: information leakage from collaborative deep learning. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 603-618).

[2] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y. and Change Loy, C., 2018. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (pp. 0-0).

[3] Smith, R., 2007, September. An overview of the Tesseract OCR engine. In *Ninth international conference on document analysis and recognition (ICDAR 2007)* (Vol. 2, pp. 629-633). IEEE.

[4] Chua, L.O. and Roska, T., 1993. The CNN paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, *40*(3), pp.147-156.

[5] Yang, J., Wright, J., Huang, T.S. and Ma, Y., 2010. Image super-resolution via sparse representation. *IEEE transactions on image processing*, *19*(11), pp.2861-2873.

[6] Lauer, F., Suen, C.Y. and Bloch, G. (2007) 'A trainable feature extractor for handwritten digit recognition', Pattern Recognition, 40(6), pp. 1816–1824. doi:10.1016/j.patcog.2006.10.011.

[7] Knerr, S., Personnaz, L. and Dreyfus, G. (1992) 'Handwritten digit recognition by neural networks with single-layer training', IEEE Transactions on Neural Networks, 3(6), pp. 962–968. doi:10.1109/72.165597.

[8] Prechelt, L., 1998. Automatic early stopping using cross validation: quantifying the criteria. *Neural networks*, *11*(4), pp.761-767.

[9] *Y. Le Cun, B. Boser, J. S. Denker, R. E. Howard, W. Habbard, L. D. Jackel, and D. Henderson. 1990. Handwritten digit recognition with a back-propagation network. Advances in neural information processing systems 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 396–404.*

[10] Chien, Y. (1974) 'Pattern classification and scene analysis', *IEEE Transactions on Automatic Control*, 19(4), pp. 462–463. doi:10.1109/tac.1974.1100577.

[11] Ferri, C., Hernández-Orallo, J. and Modroiu, R., 2009. An experimental comparison of performance measures for classification. *Pattern recognition letters*, *30*(1), pp.27-38.

[12] Bottou, L., Cortes, C., Denker, J.S., Drucker, H., Guyon, I., Jackel, L.D., LeCun, Y., Muller, U.A., Sackinger, E., Simard, P. and Vapnik, V., 1994, October. Comparison of classifier methods: a case study in handwritten digit recognition. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. 3-Conference C: Signal Processing (Cat. No. 94CH3440-5)* (Vol. 2, pp. 77-82). IEEE.

[13] Liu, C.L., Nakashima, K., Sako, H. and Fujisawa, H., 2003. Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern recognition*, *36*(10), pp.2271-2285.

[14] van Breukelen, M., Duin, R.P., Tax, D.M. and Den Hartog, J.E., 1998. Handwritten digit recognition by combined classifiers. *Kybernetika*, *34*(4), pp.381-386.

[15] Bondzulic, B.P., Pavlovic, B.Z., Petrovic, V.S. and Andric, M.S., 2016. Performance of peak signal-to-noise ratio quality assessment in video streaming with packet losses. *Electronics Letters*, *52*(6), pp.454-456.

[16] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), pp.600-612.

[17] Le Cun, Y., Jackel, L.D., Boser, B., Denker, J.S., Graf, H.P., Guyon, I., Henderson, D., Howard, R.E. and Hubbard, W., 1989. Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine*, *27*(11), pp.41-46.